УДК 621.391.6

ОБ ОЦЕНИВАНИИ ПАРАМЕТРОВ ЛИНЕЙНЫХ МОДЕЛЕЙ МНОГОМЕРНЫХ СИГНАЛОВ[1]

Е.Г. ЖИЛЯКОВ
С.П. БЕЛОВ

Белгородский государственный национальный исследовательский университет

e-mail:
Zhilyakov@bsu.edu.ru
Belov@bsu.edu.ru

В статье рассмотрены некоторые аспекты определения количественных параметров линейных моделей согласованности изменений нескольких переменных величин в предположении, что их эмпирические значения не тождественны «истинным». Такие ситуации возникают достаточно часто, например, из-за наличия погрешностей регистрации эмпирических данных, обусловленных измерительной аппаратурой. В качестве типичного примера можно указать задачу идентификации, когда определяются коэффициенты линейной модели вход – выход. Можно также отметить ситуации, когда оценки параметров, полученные на этапе обучения, используются при обработке данных, регистрируемых в иные интервалы времени, например, в задачах распознавания образов или автокомпенсации помех. В некоторых случаях речь может идти о моделях взаимовлияния (взаимосвязи) исследуемых процессов.

Исследована возможность использования в этих условиях принципа ортогонального проектирования эмпирических значений на гиперплоскость, определяемую совокупностью анализируемых «истинных» величин. Получены основные вычислительные соотношения для оценивания параметров линейных моделей согласованности.

Ключевые слова: линейные модели согласованности (взаимосвязи), оценивание параметров, ортогональное проектирование

Введение и постановка задачи

В рамках данной работы линейной моделью согласованности (взаимосвязи) некоторой совокупности процессов $Z_i=\{z_{ki}\}, i=1,..,M; k=1,2,..$ называется уравнение

$$\sum_{i=1}^{M} a_i z_{ki} = 0, k=1,2,.., \tag{1}$$

где $a_i, i=1,..,M..$ - вещественные параметры, из которых, по крайней мере, два не равны нулю.

Эта модель часто используется в ситуациях, когда наблюдаемая совокупность переменных величин порождается одним источником, либо эволюционирует под его сильным влиянием. В качестве реальных соответствующих ситуаций можно указать проявления общего для различных радиоканалов источника помех [1] или согласованное поведение цен на акции различных компаний, определяемого объемом денежного оборота на фондовом рынке. В некоторых случаях согласованность изменений процессов может быть обусловлена наличием взаимного влияния переменных, например, стоимости акций компаний, деятельность которых связана с одной и той же сферой производства.

В некотором смысле канонической формой модели (1) является [2]

$$y_k = \sum_{i=1}^{M} b_i x_{ki}, k=1,2,.. \tag{2}$$

При этом переменная слева от знака равенства по каким – либо причинам выделяется в качестве зависимой от переменных в правой части. Например, переменные в правой части могут представлять собой специально организуемые при идентификации воздействия.

[1] Исследования выполнены при поддержке гранта РФФИ № 12-07-00514-а, Государственного задания НИУ «БелГУ» на 2014 год (код проекта № 358)

Таким способом отмечается наличие причинно – следственных связей, когда изменения одной переменной обусловлены изменениями нескольких других. В частности такие ситуации возникают в задачах идентификации систем.

С некоторыми оговорками сюда также можно отнести однородные разностные уравнения, которым удовлетворяют последовательности значений одной и той же переменной

$$y_k = \sum_{i=1}^{M} c_i y_{k-i}, k = 1,2,.. \quad (3)$$

В частности таким уравнениям подчиняются автокорреляционные функции процессов авторегрессии [3], а также последовательности вида [4]

$$y_k = \sum_{i=1}^{p} b_i \cos(\omega_i k), k = 1,2,.. \quad (4)$$

где $\omega_i, i = 1,..,p$ - нормированные круговые частоты, измеряемые в радианах и, предполагается выполнение равенства

$$M = 2p. \quad (5)$$

Отметим, что соотношение (4) описывает реакцию некоторой системы в зависимости от частот воздействий, например, при определении коэффициентов усиления/ослабления сигналов в каналах передачи информации.

При известных параметрах (коэффициентах) в правых частях соотношения (2)-(3) позволяют решать прямую задачу – вычисление одних процессов по известным значениям других. Гораздо большие трудности представляет обратная задача – определение параметров уравнений по зарегистрированным значениям анализируемой совокупности процессов. Наиболее универсальный подход к её решению базируется на принципе (методе) наименьших квадратов (МНК), когда в качестве наиболее приемлемого принимается набор параметров, являющийся решением вариационной задачи

$$F = \sum_{k=1}^{N} (y_k - \sum_{i=1}^{M} b_i x_{ki})^2 = \min \sum_{k=1}^{N} (y_k - \sum_{i=1}^{M} d_i x_{ki})^2, \forall \vec{d} = (d_1,..,d_M)^T \in R^M \quad (6)$$

Здесь и в дальнейшем предполагается, что анализу подвергаются наборы по N значений каждого процесса, символ T в обозначении векторов означает транспонирование, R^M - пространство вещественнозначных векторов соответствующей размерности.

Предполагая справедливым неравенство

$$N \geq M \quad (7)$$

нетрудно получить уравнение для оптимального в смысле условия (6) вектора параметров

$$A\vec{b} = \vec{r}, \quad (8)$$

где $A = \{a_{mn}\}, m,n = 1,..,M$; $\vec{r} = (r_1,..,r_M)^T$;

$$a_{mn} = \sum_{k=1}^{N} x_{kn} x_{kn}, \quad (9)$$

$$r_m = \sum_{k=1}^{N} y_k x_{km}, m = 1,..,M . \quad (10)$$

Если процессы в правой части (2) являются линейно независимыми и, следовательно, матрица A является неособенной, то уравнение (8) разрешимо

$$\vec{b} = A^{-1}\vec{r} , \quad (11)$$

а значение функционала (6) будет равно

$$F = \| \vec{y} \|^2 - \vec{r}^T A^{-1}\vec{r} , \quad (12)$$

где символ $\| \bullet \|^2$ означает квадрат евклидовой нормы вектора (или матрицы) [5].

Легко понять, что если исходные переменные подчиняются соотношению (2), то функционал (6) будет равен нулю а, следовательно (11) дает точные значения искомых параметров.

Другой характерный признак «точно» согласованных процессов заключается в специфическом свойстве инвариантности параметров вида (11) к изменениям выделяемого в качестве левой части в (2) зависимого от других (то есть выделению процесса с единичным коэффициентом в (1)). Например, при преобразовании (2) к виду

$$x_{k1} = c_1 y_k - \sum_{i=2}^{M} c_i x_{ki}, k = 1,2,.. \quad (13)$$

входящие сюда параметры соотносятся с получаемыми из (11) следующим образом:

$$c_1 = 1/b_1; \; c_i = b_i / b_1, \;\; i = 2,..,M \quad (14)$$

Отметим также, что найденные таким способом параметры инвариантны также в том смысле, что уравнение (2) будет выполняться для значений строго согласованных процессов, регистрируемых в другие моменты времени. Именно это предположение неявно используется в задачах распознавания образов, автокомпенсации искусственно создаваемых помех, прогнозе и т.п.

Вместе с тем наличие строгой согласованности в смысле (1) является скорее исключением, чем правилом особенно в случае использования отягощенных ошибками регистрации эмпирических данных. При этом соотношения вида (14) уже не будут выполняться. Кроме того, при использовании полученных параметров для эмпирических данных, зарегистрированных в иное время, погрешность выполнения условия (2) может существенно отличаться от вычисленного ранее на основе (12) значения.

Поэтому при оценивании параметров уравнений вида (2) представляется более адекватным использовать предположение о том, что не только его левая часть, но и значения процессов в правой части известны неточно. В качестве основной модели наличия неопределенностей естественно (для МНК) использовать аддитивные смеси вида

$$\mathrm{y_k} = \sum_{i=1}^{M} p_{\mathrm{i}} s_{ki} + \varepsilon_k^y , \quad (15)$$

$$\mathrm{x_{ki}} = s_{ki} + \varepsilon_{ki}^x, i = 1,..,M . \quad (16)$$

Здесь символом ε с индексами обозначены отличия значений соответствующих гипотетических процессов от регистрируемых.

Отметим, что в общем случае по отдельности элементы правых частей в последних соотношениях являются неизвестными, включая вероятностные свойства погрешностей.

Поэтому целесообразно снова применить некоторый вариационный принцип, позволяющий получить вычислительные соотношения для оценок, как искомых параметров, так и гипотетических процессов. Очевидно, что в последнем случае речь идет о варианте фильтрации.

В рамках данной работы используется принцип проектирования векторов

$$\vec{u}_k = (x_{k1},..,x_{kM}, y_k)^T \tag{17}$$

на линеалы [6] вида

$$\mathrm{V_k} = \sum_{i=1}^{M} d_{\mathrm{i}} s_{ki} \tag{18}$$

Но основе вычислительных экспериментов исследуются перечисленные выше задачи оценивания параметров согласованности.

Вычислительные соотношения

Пусть $\vec{d} = (d_1..,d_M)^T \in R^M$ – некоторый фиксированный набор параметров. Положим

$$G = \sum_{k=1}^{N} G_k, \tag{19}$$

где

$$G_k = \sum_{i=1}^{M} (x_{ki} - s_{ki})^2 + (y_k - \sum_{i=1}^{M} d_i s_{ri})^2. \tag{20}$$

Чтобы вектор

$$\vec{s}_k = (s_{k1},..,s_{kM})^T \tag{21}$$

был проекцией вектора вида (17) на линеал (18) необходимо и достаточно выполнения системы уравнений

$$\partial D_k / \partial s_{kr} = 0, r = 1,..,M. \tag{22}$$

В результате дифференцирования (20) получаем уравнение

$$B\vec{s}_k = \vec{x}_k + y_k \vec{d}, \tag{23}$$

где $\vec{x}_k = (x_{k1},..,x_{kM})^T$;

$$B = I + \vec{d}\vec{d}^T, I = diag(1,..,1). \tag{24}$$

Нетрудно показать справедливость соотношения для обратной к (24) матрицы

$$B^{-1} = I - h\vec{d}\vec{d}^{T}, \tag{25}$$

где

$$h = 1/(1 + \| \vec{d} \|^2). \tag{26}$$

Поэтому решение уравнения (23) имеет вид

$$\vec{s}_k = \vec{x}_k + h(y_k - (\vec{x}_k, \vec{d}))\vec{d}, \tag{27}$$

где символ $(.,.)$ означает скалярное произведение в евклидовом пространстве векторов.

Отметим, что соотношение (27) при заданном векторе $\vec{d}$ можно интерпретировать как схему фильтрации значений вектора $\vec{x}_k$, при этом, если имеет место равенство $y_k = (\vec{x}_k, \vec{d})$, то соотношение (27) дает

$$\vec{s}_k = \vec{x}_k, \tag{28}$$

то есть этот вектор не искажается. В свою очередь в качестве отфильтрованных значений последовательности y_k из (27) получаем

$$\hat{y}_k = (\vec{s}_k, \vec{d}) = h(\| \vec{d} \|^2 \; y_k + (\vec{x}_k, \vec{d})). \tag{29}$$

Ясно, что при выполнении оговоренного выше условия наряду с (28) левая часть соотношения (29) будет равна значению исходной последовательности.

Подстановка (27) в (20) дает

$$G_k = h(y_k - (\vec{x}_k, \vec{d}))^2, \tag{30}$$

а соответственно (19) преобразуется к виду

$$G = h\sum_{k=1}^{N} (y_k - (\vec{x}_k, \vec{d}))^2. \tag{31}$$

Дальнейшее связано с поиском вектора параметров, удовлетворяющего вариационному условию

$$W = \sum_{k=1}^{N} (y_k - (\vec{x}_k, \vec{p}))^2 /(1 + \| \vec{p} \|^2) = \min \sum_{k=1}^{N} (y_k - (\vec{x}_k, \vec{d}))^2 /(1 + \| \vec{d} \|^2), \forall \vec{d} \in R^N \tag{32}$$

то есть к минимизации функционала

$$D(\vec{d}) = \sum_{k=1}^{N} (y_k - (\vec{x}_k, \vec{d}))^2 /(1 + \| \vec{d} \|^2), \vec{d} \in R^N. \tag{33}$$

Отметим, что выполнение замен вида

$$d_i = d_r c_i, i = 1,..,M; y_k \Leftrightarrow x_{kr} \tag{34}$$

приводит к равенству

$$D(\vec{d}) = D(\vec{c}), c_r = 1/d_r, c_i = d_i/d_r, i \neq r. \tag{35}$$

Иными словами выполняется упоминаемое выше свойство инвариантности (симметрии относительно переобозначений вида (34)).

Положим

$$\vec{d} = \vec{b} + \vec{w}, \tag{36}$$

где первый вектор в правой части определяется из соотношения (11). Тогда с учетом (12) (33) нетрудно преобразовать к виду

$$D(\vec{d}) = D(\vec{w}) = (F + \vec{w}^T A\vec{w})/(1 + \|\vec{b} + \vec{w}\|^2)., \vec{w} \in R^N. \tag{37}$$

Таким образом, задача (32) сводится к поиску соответствующего вектора $\vec{w}$.

Пусть Q и $L = diaq(\lambda_1,..,\lambda_M)$ - матрицы собственных векторов и соответствующих им собственных чисел матрицы *A*, так что выполняется условие

$$AQ = QL, \tag{38}$$

причем собственные числа упорядочены по убывания

$$\lambda_k > \lambda_{k+1} > 0, k = 1,..,M-1. \tag{39}$$

Полагая

$$\vec{b} = Q\vec{\alpha}; \vec{w} = Q\vec{\beta}, \tag{40}$$

функционал (37) нетрудно преобразовать к более простому для анализа виду

$$D(\vec{\beta}) = (F + \vec{\beta}^T L\vec{\beta})/(1 + \|\vec{\alpha} + \vec{\beta}\|^2), \vec{\beta} \in R^N. \tag{41}$$

Отсюда после дифференцирования по компонентам $\vec{\beta}$ получаем соотношение для градиента минимизируемого функционала

$$\vec{g}(\vec{\beta}) = -[(F + \vec{\beta}^T L\vec{\beta})\vec{\alpha}/(1 + \|\vec{\alpha} + \vec{\beta}\|^2) - L\vec{\beta}]/(1 + \|\vec{\alpha} + \vec{\beta}\|^2) \tag{42}$$

Это соотношение определяет следующую схему градиентного спуска при минимизации функционала (41)

$$\vec{\beta}_k = \vec{\beta}_{k-1} - c_k\vec{g}(\vec{\beta}_{k-1}), 0 < \mathrm{c}_{\mathrm{k-1}} << 1;\ \vec{\beta}_0 = 0 \tag{43}$$

Очевидно, что

$$\vec{g}(0) = -F\vec{\alpha}/(1 + \|\vec{\alpha}\|^2)^2 = -D(0)\vec{\alpha}/(1 + \|\vec{\alpha}\|^2 \tag{44}$$

поэтому

$$\vec{\beta}_1 = c_1\vec{\alpha}, 0 < \mathrm{c}_1 << 1, \tag{45}$$

и (42) дает

$$\vec{g}(\vec{\beta}_1) = -H_1\vec{\alpha} \tag{46}$$

где *H* – диагональная матрица $H_1 = diag(h_{11},..,h_{M1})$

$$h_{k1} = [(F + c_1^2\alpha^T L\vec{\alpha})/(1 + (1 + c_1)^2 \|\vec{\alpha}\|^2) - c_1\lambda_k]/(1 + (1 + c_1)^2 \|\vec{\alpha}\|^2 .) \tag{47}$$

В свою очередь схема (43) дает представление

$$\vec{\beta}_2 = (c_1 I + c_2 H_1)\vec{\alpha} = B_2\vec{\alpha}\,, \tag{48}$$

подстановка которого в (42) позволяет получить соотношение

$$\vec{g}(\vec{\beta}_2) = -H_2\vec{\alpha}\,, \tag{49}$$

где элементы матрицы $H_2 = diag(h_{12},..,h_{M2})$ легко определяются после несложных преобразований.

Легко понять, что на каждом шаге последовательности (43) будут получаться диагональные матрицы, то есть будет получаться последовательность векторов

$$\vec{\beta}_k = B_k\vec{\alpha},\; B_k = c_1 I + \sum_{i=1}^{k-1} c_{i+1}H_i\,. \tag{50}$$

Эта последовательность будет сходиться к решению задачи вида (32).

Список литературы

1. Ширман Я.Д., Манжос В.Н. Теория и техника обработки радиолокационной информации. М.: Радио и связь, 1981.
2. Большаков А.А., Каримов Р.Н. Методы обработки многомерных данных и временных рядов М.: Горячая линия-Телеком, 2007.
3. Андерсон Т. Статистический анализ временных рядов. М.: Мир, 1975.
4. Марпл С.Л-мл. Цифровой спектральный анализ и его приложения. М.: Мир, 1990.
5. Гантмахер Ф.Р. Теория матриц. М.:Наука, 1967.
6. Ректорис К. Вариационные методы в математической физике и технике. М.: Мир, 1985.

ABOUT THE ESTIMATION OF PARAMETERS OF LINEAR MODELS OF MULTIDIMENSIONAL SIGNALS

E.G. ZHILYAKOV
S.P. BELOV

Belgorod State National Research University,

e-mail:
Zhilyakov@bsu.edu.ru
Belov@bsu.edu.ru

In the article some aspects of the definition of the quantitative parameters of linear models of consistency changes of several variables on the assumption that their empirical values are not identical to "true". Such situations arise, for example, because of the presence of errors registration of empirical data, due to the measuring equipment. As a typical example, you can specify the task of identification when determining the coefficients of the linear model of input – output. You may also notice a situation when valuation parameters obtained during the training phase, are used in the processing of data recorded in different time intervals, for example in the problems of pattern recognition or self-compensation interference. In some cases we can speak about models of the interaction (relationship) of the studied processes.

Explore the feasibility of using in these conditions the principle of orthogonal projecting of empirical values on the hyperplane defined by a set of analyzed the "true" values. Received basic computational formulas for estimating the parameters of linear models of consistency.

Key words: linear model of consistency (the relationship), parameter estimation, orthogonal design